

Случайные процессы. Прикладной поток.

Практическое задание 10.

Методы обнаружения разладок.



Правила:

- Выполненную работу нужно отправить на почту probability.diht@yandex.ru, указав тему письма "[СП17] Фамилия Имя - Задание 10". Квадратные скобки обязательны. Вместо Фамилия Имя нужно подставить свои фамилию и имя.
- Прислать нужно ноутбук и его pdf-версию. Названия файлов должны быть такими: 10.N.ipynb и 10.N.pdf, где N — ваш номер из таблицы с оценками.
- Никакой код из данного задания при проверке запускаться не будет.

В файле *ftse100.csv* (отсюда) содержится информация о значении индекса FTSE за каждый рабочий день на момент закрытия с января 1980 года по март 2004 года.

В файле *internet-traffic.csv* (отсюда) содержится информация о величине интернет-трафика (в битах) в академической сети Великобритании с 19 ноября 2004 года до 27 января 2005 года за каждый час.

Задание:

1. Нарисуйте график временного ряда. По графику на глаз определите, какие типы разладок могут происходить. Это могут быть, например, разладки "увеличение среднего", "уменьшение дисперсии". По данным не нужно определять величину изменения характеристики распределения при разладке. Нужно только решить, какие по величине отклонения вы считаете значимыми.
2. По небольшому начальному отрезку времени, в течении которого точно нет разладки, оцените распределение P_∞ . Над данными можно совершать любые преобразования, однако нужно следить за тем, чтобы в каждый момент времени не использовались данные из будущего.
3. По временному ряду запустите статистики CUMSUM и Ширяева-Робертса для детектирования всех предполагаемых типов разладок (т.е. одновременно несколько статистик для разных распределений).
4. Обнаруживают ли статистики все имеющиеся разладки во временном ряду? Какая при этом задержка при подаче сигнала? Сколько ложных тревог?

За исследование данных по каждому файлу можно получить 5 баллов.